Retrovirology

Dynamic features of the selective pressure on the human immunodeficiency virus type 1 (HIV-1) gp120 CD4-binding site in a group of long term non progressor (LTNP) subjects

Filippo Canducci*[1], Maria Chiara Marinozzi[1], Michela Sampaolo[1], Stefano Berrè[1], Patrizia Bagnarelli[2], Massimo Degano[3], Giulia Gallotta[4], Benedetta Mazzi[5], Philippe Lemey[6], Roberto Burioni[1] and Massimo Clementi[1]

Address: [1]Laboratorio di Microbiologia e Virologa, Università Vita-Salute San Raffaele, Milan, Italy, [2]Istituto di Microbiologia, Università Politecnica delle Marche, Ancona, Italy, [3]Unità di Biocristallografia, Istituto Scientifico San Raffaele, Milan, Italy, [4]Dipartimento di Malattie Infettive, Università Vita-Salute San Raffaele, Milan, Italy, [5]Laboratorio di Ematologia Molecolare, Istituto Scientifico San Raffaele, Milan, Italy and [6]Rega Institute, Katholieke Universiteit Leuven, Leuven, Belgium

Email: Filippo Canducci* - canducci.filippo@hsr.it; Maria Chiara Marinozzi - marinozzi.mariachiara@hsr.it; Michela Sampaolo - sampaolo.michela@hsr.it; Stefano Berrè - stefanoberre@yahoo.it; Patrizia Bagnarelli - bagnarelli@univpm.it; Massimo Degano - degano.massimo@hsr.it; Giulia Gallotta - gallotta.giulia@hsr.it; Benedetta Mazzi - mazzi.benedetta@hsr.it; Philippe Lemey - philippe.lemey@gmail.com; Roberto Burioni - burioni.roberto@hsr.it; Massimo Clementi - clementi.massimo@hsr.it

* Corresponding author

Abstract

The characteristics of intra-host human immunodeficiency virus type 1 (HIV-1) *env* evolution were evaluated in untreated HIV-1-infected subjects with different patterns of disease progression, including 2 normal progressor [NP], and 5 Long term non-progressor [LTNP] patients. High-resolution phylogenetic analysis of the C2-C5 *env* gene sequences of the replicating HIV-1 was performed in sequential samples collected over a 3–5 year period; overall, 301 HIV-1 genomic RNA sequences were amplified from plasma samples, cloned, sequenced and analyzed. Firstly, the evolutionary rate was calculated separately in the 3 codon positions. In all LTNPs, the 3rd codon mutation rate was equal or even lower than that observed at the 1st and 2nd positions (p = 0.016), thus suggesting strong ongoing positive selection. A Bayesian approach and a maximum-likelihood (ML) method were used to estimate the rate of virus evolution within each subject and to detect positively selected sites respectively. A great number of N-linked glycosylation sites under positive selection were identified in both NP and LTNP subjects. Viral sequences from 4 of the 5 LTNPs showed extensive positive selective pressure on the CD4-binding site (CD4bs). In addition, localized pressure in the area of the IgG-b12 epitope, a broad neutralizing human monoclonal antibody targeting the CD4bs, was documented in one LTNP subject, using a graphic colour grade 3-dimensional visualization. Overall, the data shown here documenting high selective pressure on the HIV-1 CD4bs of a group of LTNP subjects offers important insights for planning novel strategies for the immune control of HIV-1 infection.

Background

Virus-host relationships in human immunodeficiency type 1 virus (HIV-1) infection are characterized by a great complexity. The virus is strictly dependent on the host cell for replication, but it is constantly exposed to the immune response of the infected host. Although the innate and adaptive immune responses restrict HIV-1 replication after primary infection [1-3], efficient control of virus replication and consequent stable levels of CD4+ T-cells are observed only in a minority of patients designated long-term non progressors (LTNPs). In LTNPs virus replication is limited, suggesting that HIV-1 variants are less fit than those detectable in normal or rapid progressors in this subgroup of infected persons [4]-. Since in the absence of anti-retroviral therapy (ART), the HIV-1 replication capacity (RC) is largely related to the efficiency of viral entry [5,6]-, the selective pressure exerted either by CTL or neutralizing antibodies can account for particular evolutionary patterns in the *env* gene in LTNPs [7-10].

HIV-1 evades the immune response of the host using different mechanisms, including steric occlusion, conformational masking of critical parts of the protein, and insertions or deletions in variable loops [2,11]. Additionally, the vast majority of antibodies directed against the viral envelope recognize non-neutralizing epitopes of the glycoprotein monomers, thus probably being ineffectual against the trimeric functional complex [6,12]. Furthermore, a shifting "glycan shield" has been shown to protect the virus from neutralization by monoclonal antibodies [13-16]. Finally, many envelope surface elements are believed to serve as a decoy for the host immune system, being largely tolerant to variation with no effect on virus RC [17]. However, conserved *env* regions have been described and they are generally associated with functional properties, including virus binding to receptors and co-receptors. In particular, the CD4 binding-site (CD4bs) is believed to be a highly conserved region exposed to the solvent for ligand binding [18]-. In LTNPs, control of virus replication seems to correlate with the presence of antibodies against this critical domain, and sera from these patients show broad cross-neutralizing responses against primary HIV-1 isolates, mainly due to antibodies against this epitope [19-22].

In the past few years, a growing body of studies has investigated the HIV-1 *env* gene evolution in order to evaluate its role during the natural course of infection [19,23-27], and to identify the crucial characteristics of active and passive immunization strategies [15,18,20,28-30]-. Positively selected sites have frequently been observed within the C2-V5 region of the viral surface glycoprotein in samples from recently and chronically infected patients [1,9,10,23,24,26,27,31,32]. In the present study, a high-resolution phylogenetic analysis of partial *env* gene nucleotide sequences (C2-C5 region) was performed using samples collected over a period of 3–5 years from 7 HIV-1 infected, untreated, asymptomatic patients with different patterns of disease progression. The aim of this study was to identify conformational epitopes and sites of the viral protein surface with specific patterns of virus evolution in LTNPs.

Results

HIV-1 evolutionary rate in normal progressors and in long-term non progressor patients

Virus evolutionary rate (substitutions/site/year) within each patient was estimated separately for the first + second ($\mu^{1st+2nd}$) and third codon position (μ^{3rd}) separately (Figure 1). The average viral mutation rate among all patients was estimated to be around 2.34E-02 mutations/site/year. In patients A, B (normal progressors; NP), the average mutation rate (μ) was significantly higher at the third position compared to that of the first and second positions ($\mu3^{rd}$ compared to $\mu^{1st+2nd}$). In all LTNPs, the third codon mutation rate was estimated to be lower or almost equal to that inferred for the other codon positions ($\mu3^{rd}$ compared to $\mu^{1st+2nd}$). This difference was found to be statistically significant when LTNP and NP results were compared with the Student t-Test (p = 0,016).

Maximum likelihood analysis of positive selection on non recombinant data sets

We compared the fit of two sets of nested site-specific models to the data (including a neutral model that is restricted to purifying selection and an alternative model that also allows for positive selection): Model 1a vs. Model 2a and Model 7 vs. Model 8. To assess whether allowing codons to evolve under positive selection gives a significantly better fit to the data, the log likelihood values obtained for each pair of nested models were compared using the Likelihood Ratio Test (LRT) (Additional file 1). In all cases Model 2a and Model 8 were significantly favoured over Model 2a and Model 7 respectively (P < 0.001), and the empirical Bayes approach identified several positively selected sites.

Site specific dN/dS values for each patient and the entropy value for each position along the sequence were calculated (data not shown). Subsequently, a color-grade 3-dimensional visualization of the dN/dS score (the posterior mean value derived from the Empirical Bayes approach using Model M8) was generated (Figure 2 and 3). Using Model 8, the following numbers of sites with a dN/dS ratio higher than 1 were observed: patient A: 24; patient B: 33, patient C: 53; patient D: 45; patient E: 45; patient F: 81 patient G: 52. The following number of sites with dN/dS > 2 were observed: patient A: 15; patient B: 23, patient C: 27; patient D: 36; patient E: 33; patient F: 56 patient G: 34. The following numbers of sites with a dN/
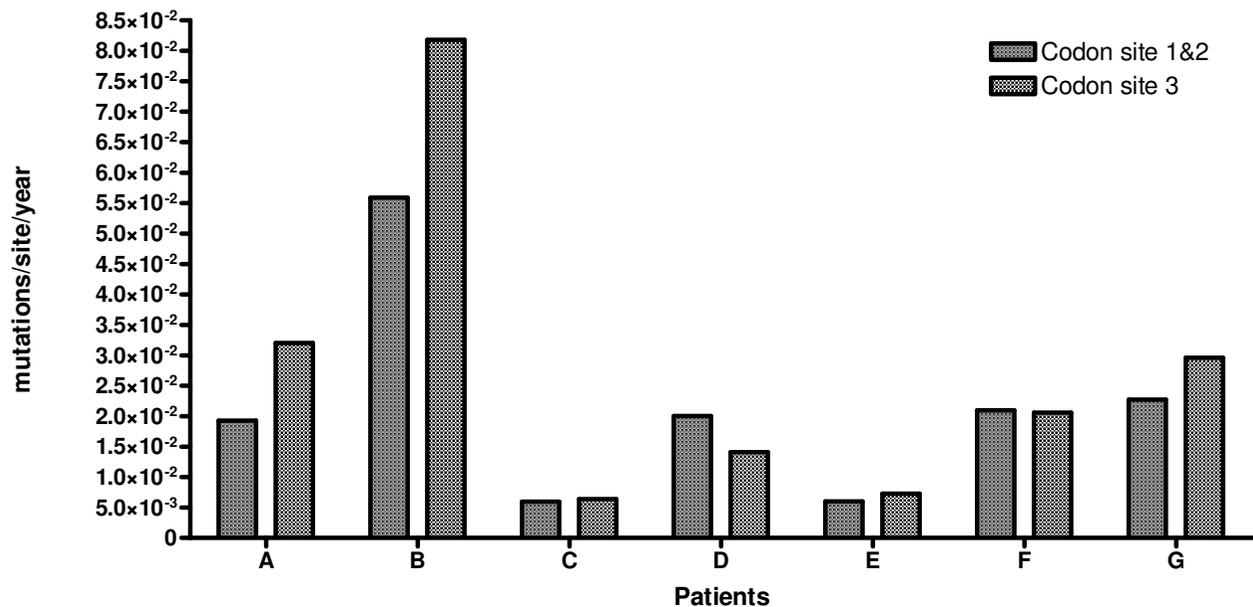
Figure 1
Site specific mutation rate. Virus mutation rate (mutations/site/year) within each patient. For each patient the mutation rate for each codon position was estimated.

dS ratio higher than 3 were observed: patient A: 13; patient B: 0, patient C: 19; patient D: 25; patient E: 23; patient F: 42; patient G: 17.

The following number of sites with a posterior probability of being under positive selection > 95% and > 99%, respectively, were identified: patient A: 6 and 4; patient B: 7 and 1; patient C: 8 and 3; patient D: 10 and 7; patient E: 9 and 5; patient F: 23 and 11; patient G: 8 and 2. Selective constraints appear to act along all the proteic sequence in all patients. In all patients, positively selected sites appeared to be unevenly distributed. In particular the majority of sites were located in C3 and in V4, where many N-linked glycosylation sites are known to be present and used to protect from antibody mediated neutralization [30].

To examine the molecular footprint of deleterious mutational load on within-host evolution, and its putative impact on the identification on positively selected sites, we tested for differences in selective pressure among internal and external branches in each patient. dN/dS estimates were almost always higher on external branches compared to internal branches, but only for three patients this was statistically supported by the LRT model comparison (see Additional file 2). When the internal-external differences were tested on the data combined for all patients, however, a higher dN/dS on external branches

(0.46 for internal vs 0.78 for external) was strongly supported by the LRT (< 0.001). This analysis confirms that external branches are subject to deleterious load, which might result in an elevated dN/dS ratio for these branches [33]. When we inferred the sites under selection only for the internal branches using the Fixed Effects Likelihood (FEL), several of the sites identified using the previous models were confirmed to be under positive selection (Figure 4).

For the 5 patients for which the HLA typing was obtained (see below), the majority of positively selected sites were localized outside the known HLA class I linear epitopes except for patients B, C, and E, where residues immediately next to or belonging to an HLA-A11 epitope were identified (position 339 to 350). In particular, in patient B and E residues 344Q (that is also exposed on the surface) and 346A and position 339N in patient C was inferred to be under positive selection.

3-dimesional analysis of the dN/dS score
A 3-dimensional visualization of the posterior mean dN/dS value was generated using a color grade scale. Both on the CD4 binding site and on the outer domain of the molecule the majority of sites appeared as under purifying selection (Figures 2, 3 and 5, light blue areas), especially in patients C, D, and E. In many cases, amino acids that were identified as under positive selection along the
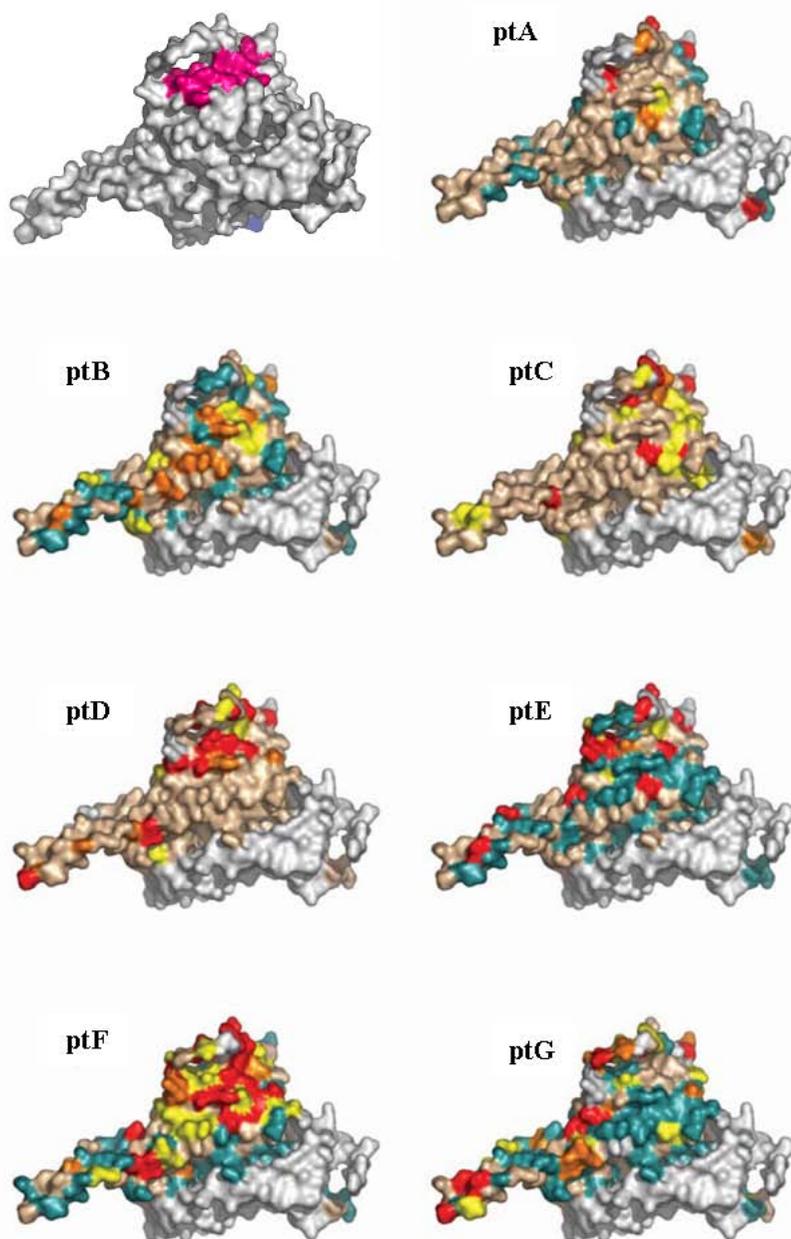
Figure 2
dN/dS score visualization on the surface of gp120 (the 'silent' face of the molecule). Visualization of the dN/dS score (the posterior mean value derived from the Empirical Bayes approach using Model M8) onto the molecular surface of gp120 (pdb code 2B4C) using a color grade scale. Sites with no data or with a dN/dS score < 0.002 are depicted in white, sites with a dN/dS score between 0.002 and 0.15 are in light blue, sites between 0.15 and 1 are in light brown, sites with a dN/dS score between 1 and 2 are yellow, sites with a dN/dS score between 2 and 3 are orange, sites with a dN/dS score > 3 are red on the surface. A gp120 molecule was added in the upper left quadrants to localize CD4 and/or IgGb12 contact residues and the C3 alpha helix. Residues that are involved only in CD4 binding are depicted in blue, residues involved in IgGb12 binding are depicted in yellow, residues that interact both with CD4 and IgGb12 are displayed in green colour (modified from Zhou et al, 2007). The alpha helix present in the C3 region is shown in magenta.

Figure 3
dN/dS score visualization on the surface of gp120 (the internal portion and the CD4 binding region). Visualization of the dN/dS score (the posterior mean value derived from the Empirical Bayes approach using Model M8) onto the molecular surface of gp120 (pdb code 2B4C) using a color grade scale. Sites with no data or with a dN/dS score < 0.002 are depicted in white, sites with a dN/dS score between 0.002 and 0.15 are in light blue, sites between 0.15 and 1 are in light brown, sites with a dN/dS score between 1 and 2 are yellow, sites with a dN/dS score between 2 and 3 are orange, sites with a dN/dS score > 3 are red on the surface. A gp120 molecule was added in the upper left quadrants to localize CD4 and/or IgGb12 contact residues and the C3 alpha helix. Residues that are involved only in CD4 binding are depicted in blue, residues involved in IgGb12 binding are depicted in yellow, residues that interact both with CD4 and IgGb12 are displayed in green colour (modified from Zhou et al, 2007). The alpha helix present in the C3 region is shown in magenta.
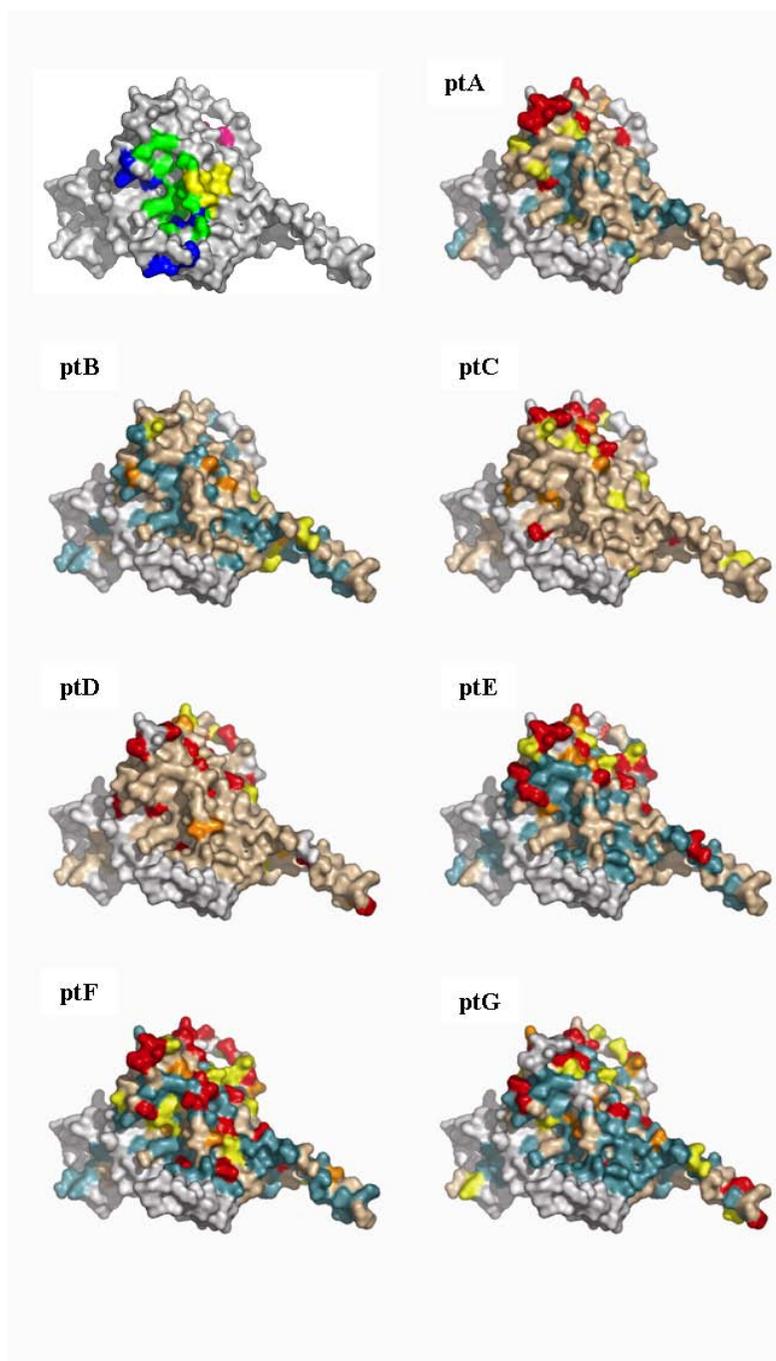
	aa position					aa position					aa position					Aa position		

pt. A

aa position		
283	S -> I	
	I -> V	
291	S -> V	
	S -> P	
	A -> S	
392	E -> S	
	E -> T	
460	N -> S	
	S -> P	
	S -> D	
462	E -> N	
	D -> N	
	R -> K	

pt. B

aa position	
279	D -> N
354	K -> G
	K -> N
444	R -> N
	N -> D
460	K -> N
	D -> K
	D -> N
	N -> T
496	V -> I

pt. C

aa position	
339	N -> H
	H -> D
	D -> N
360	V -> I
396	N -> Y
	H -> Y
	H -> P
	N -> K
399	I -> T
	T -> N
	I -> V
405	R -> K
	K -> E
	K -> N
463	D -> E
	E -> G

pt. D

aa position	
283	V -> S
	V -> I
321	G -> E
	E -> G
336	K -> T
	T -> E
	T -> A
	K -> R
	K -> E
	K -> Q
354	R -> K
	K -> E
	R -> G
360	A -> V
	V -> F
399	E -> K
	K -> T
	T -> S
	S -> K
	K -> E
	K -> N
404	V -> E
	E -> R

pt. E

aa position	
279	N -> D
	D -> N
281	A -> V
	V -> I
335	I -> R
	R -> G
	G -> E
	E -> G
	I -> G
350	R -> E
	E -> R
360	V -> A
	A -> V
	V -> I
362	S -> N
	N -> K
	N -> S
	N -> T
397	D -> Y
	Y -> N
	N -> G
	D -> V
	V -> F
405	G -> K
454	I -> L
	L -> I
460	I -> N
	I -> K
	I -> T
461	R -> T
	R -> E
462	T -> E
	T -> N
	N -> D
463	E -> N
	N -> S
	S -> N
	E -> G
	E -> Q
	E -> K
	G -> E

pt. F

aa position	
394	T -> S
	S -> T
410	S -> L
	L -> Q
	L -> E
461	T -> S
	S -> G
	G -> N
462	M -> T
	T -> E
	M -> L
500	R -> K
	K -> R

pt. G

aa position	
360	V -> A
396	D -> N
	D -> K
	K -> R
	K -> N
398	N -> V
	V -> D
401	N -> G
	G -> W
406	N -> D
	D -> I
	I -> N

Figure 4
Positively selected sites identified along internal branches. Amino acid (aa) positions are indicated according to HXB2 sequence.
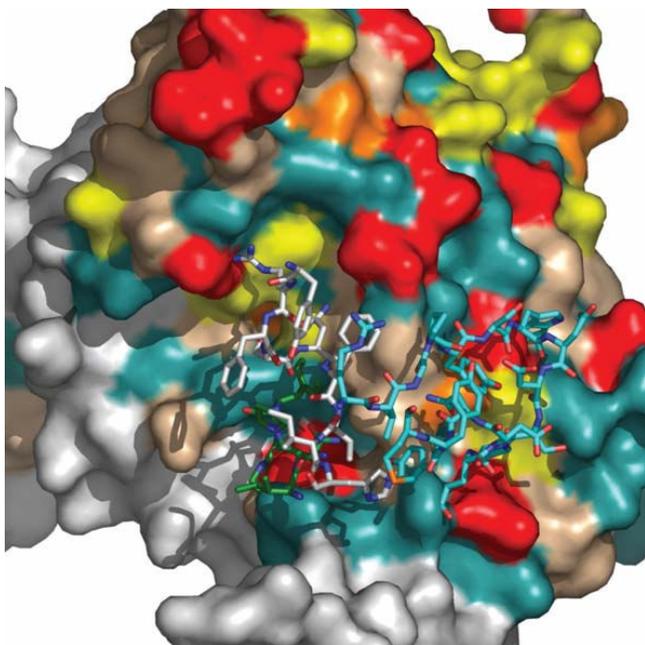
Figure 5
dN/dS score visualization on the surface of gp120 (a close-up view of the interaction site between gp120 of patient F and the IgGb12 heavy chain (pdb code NY7)). Visualization of the dN/dS score (the posterior mean value derived from the Empirical Bayes approach using Model M8) onto the molecular surface of gp120 (pdb code 2B4C) using a color grade scale. Sites with no data or with a dN/dS score < 0.002 are depicted in white, sites with a dN/dS score between 0.002 and 0.15 are in light blue, sites between 0.15 and 1 are in light brown, sites with a dN/dS score between 1 and 2 are yellow, sites with a dN/dS score between 2 and 3 are orange, sites with a dN/dS score > 3 are red on the surface. Residues that are involved only in CD4 binding are depicted in blue, residues involved in IgGb12 binding are depicted in yellow, residues that interact both with CD4 and IgGb12 are displayed in green colour (modified from Zhou et al, 2007). The alpha helix present in the C3 region is shown in magenta. The carbon atoms of CDR1, CDR2 and CDR3 are coloured white, green and cyan respectively. The amino acid residues are shown as sticks. Of note, the binding region of the broadly neutralizing antibody overlaps the positively selected sites in the patient G derived structure.

gp120 linear sequence, defined clusters on the surface, suggesting their role in conformational epitopes presented on exposed antigenic areas. In all patients a high level of variation was observed in the C3 region, where an α-helix (position 335 to 350) is located and exposed on one side to the solvent and can be recognized by humoral immune defences. On the outer domain of gp120, many clusters were identified in all patients, but with a different distribution. A conformational epitope was identified in

patient D, which was defined by Lys337, Ser334, Ala336, Asn339, Asn340 and Gln344. In patient F, a linear epitope in the C3 region that is exposed on the surface was identified and formed by Lys362, Glu363, Ser364 and Ser365. Another wide site of positive selection appeared to be formed by Glu269, Asn289, Ser291, Lys337, Gln340, Lys343, Gln344, and located on the outer surface. In patient G, the exposed surface harboured only two residues under positive selection: Ile371 and Gly471, which cluster together on the 3-D structure.

All patients had positively selected sites in the V3 region, specifically patient F (5 sites with a dN/dS > 1 located both on the tip and at its base). In all patients, no sites were identified among known CD4 induced epitopes.

Analysis of the CD4 binding site
Positively selected sites were identified in the CD4 binding region in patients C, D, E and F, but not in patients A and B, where almost all positively selected sites were located on the outer surface or on the α-helix in the C3 region. In all patients except patient B, Thr283, located in the CD4 binding region (though not directly in contact with it), was inferred to be under positive selection. In patients C and D, distinct sites were under positive selection in this area. Arg476 in patient C, and Thr283 and Asp368 in patient D, were under positive selection and potentially involved in direct receptor binding. A more clearly delimited constraint seems to act on patients E, F and G. In particular, a conformational epitope appeared to be present in patient E and G and formed by Thr278, Asp279 and Ala 281. In patient F, a complex and large area located partially within the CD4 binding site and in a usually highly conserved region immediately next to it was observed to be under positive selection. This region includes Ala281, Trp427, Glu460, Ser461, Glu462 and Leu452 and Leu453. When the IgGb12 heavy chain CDRs structures were superimposed on patient G-derived gp120 3-dimentional visualization, a high number of positively selected sites identified in this patient coincided with residues recognized by this broad neutralizing antibody on the gp120 surface [34].

Identification of rare mutations
When the amino acid entropy of positively selected sites was studied, the majority of substitutions observed for all patients were between residues present in that same position with a high frequency in the 500 database sequence alignment. Nevertheless, in some patients, rare substitutions seem to have been selected, including E269D, N339H, N339D, N340D, N340K, T341A, N343Q, N343E, A346F, A346Y, T394A, T394I, R476K, R476M. Amino acid frequencies in those positions in the 500 sequence database alignment and how these sites evolved during the observation period are shown in Table 1.

HLA typing
A low- or high-resolution HLA typing was also performed for patient A to E. HLA typing was not possible for patients F and G. Results of HLA typing are shown in Additional file 3.

Discussion
In the present study, a high-resolution phylogenetic analysis of the gp120 envelope glycoprotein evolution was performed in HIV-1 infected patients with a different pattern of disease progression. All patients under study had never been treated for HIV-1 infection, leaving the host immune system as the only selective force acting on virus evolution and quasispecies selection. Firstly, an analysis was performed to identify putative recombinants. Recombination may occur frequently in vivo in HIV-1 evolution, and artificial chimeric sequences due to PCR crossovers can significantly affect phylogenetic analysis. The PHI test based on the refined incompatibility score was used to overcome this bias with our data set [35]. When recombinant sequences were excluded (about 15%, see materials and methods) from the analysis, the number of sites with a dN/dS value > 1 was reduced in some of the patients. Nevertheless, the number of positively selected sites identified with a Bayesian posterior probability > 0.95 in our datasets was not significantly affected. The best fitting model of evolution was chosen in the phylogenetic reconstruction, and maximum likelihood methods were used to fit codon models of evolution for all patients, to identify positively selected sites, and Bayesian inference was used to estimate virus evolutionary rates. In addition, an HLA typing and a color-grade 3-dimensional visualization of the dN/dS score were used.

Finally, since external branches are subject to substitutions as well as mutational load, which involves random mutations and therefore potentially many nonsynonymous substitutions, we inferred the sites under selection for the internal branches only, using the Fixed Effects Likelihood (FEL) approach [36]. This analysis infers dN and dS for each site and also tests whether dN = dS or not for the sites [36]. All the sites identified with the FEL approach were also identified with the previous methods, further confirming the possibility of identifying sites showing diversifying selection when sequential time points are considered even using cloned sequences. A multiple-step analysis was in fact necessary in the present study to address correctly the evolution of a large portion of the HIV-1 *env* gene, since a high background is expected when the dN/dS score/site is performed in highly variable viral populations under continuous positive selection. In these cases, only sites with high dN/dS ratio and confirmed by Bayesian posterior probability should be taken into consideration [32,37,38]-.

In order to highlight the effect of positive selection on virus evolution, the evolutionary rate was calculated separately in the three codon positions. In the third codon position, mutations are silent in about 70% of all possibly occurring nucleotide changes, and if no selective constraints act on the virus, evolution occurs at a faster rate compared to the first and second codon positions. In all LTNPs, the third codon mutation rate is equal to or lower than that compared to the averaged 1st and 2nd position (p = 0.016), thus being compatible with positive selection [39-41].

The impact of HLA-associated selection pressure on viral evolution has recently been demonstrated at the population level [42-50]. No HLA B57 associated positively selected sites were identified in our patients, but a potential HLA A11 associated epitope was present in patients B, C, and E. Within this epitope, the position 346 exhibited a high dN/dS ratio in all three patients.

Although positive selection was evident in the replicating virus from all subjects, differences were observed between NPs and LTNPs. In subjects A and B (NPs) selective constraints are less intense, in terms of dN/dS score calculated even for the highly selected hotspots (Figure 2 and 3), and are limited to the external surface of the crystal and to the α-helix in the C3 region. These sites and the V3 loop appear to be targets for the immune response in all patients, with a single exception (patient A). This observation is apparently in contrast with the results obtained by other studies, where the C3 alpha helix was observed to be under positive selection for clade C envelopes and only modestly for clade B [27,51]. Although we cannot exclude that differences in the intensity of the immune response against different HIV-1 subtypes exist at these levels, the previous analyses were based on cross-sectional C-clade and B-clade sequence datasets downloaded from HIV-1 databases, thus not reflecting the intra-patient evolutionary dynamics and the heterogeneity of host immune responses during the different phases of HIV-1 infection (or the different patterns of disease progression observed). Other studies analyzed the sequence evolution in infected individuals and showed that the C3 region, including the externally accessible residues, is under strong positive selection both in clade B [24-26] and in HIV-1 subtype C infections [23]. These results may be of particular interest since this antigenic portion of the gp120 molecule has been considered in the development of candidate vaccines [52-56]-.

Many N-linked glycosylation sites were identified to be under positive selection and exposed on the surface in the group of LTNPs and in the 2 NP subjects. In particular N442, R444 and S446, N295, N332, N340, N339 were identified as being potentially involved in the glycan

Table 1: Evolution of positively selected amino acids that were rarely found in the 500 sequences database.

AA pos	Frequency in the database	A			B			C			D			E			F			G		
		I	II	III	I	II	III	I	II	III	I	II	III	I	II	III	I	II	III	I	II	III
269	Glu 76.8% **Asp 8.1%**	E			E			E			**D** **15/15**	**D** **15/15**	**D** **10/13** E 3/13	E			E			**D** **19/19**	**D 9/14** E 5/14	**D** **9/10** E 1/10
339	Asn 78.5% **Asp 4.3%** **His 0.5%**	N			N			N 8/13 **D** **4/13** **H** **1/13**	N 14/15 **H** **1/15**	N 8/14 **H** **6/14**	V			N			K	K	N	N		
340	Lys 32.5% Asn 31.3% **Asp 7.2%**	E			**D** **16/16**	**D** **3/15** N 12/15	N 10/10	**D**			K			N			N			N		
341	Thr 91.4% **Ala 5.7%**	T			T			T			T 13/15 **A** **2/15**	T 5/12 **A** **7/12**	T 1/3 **A** **12/13**	T			T			T		
343	**Asn 2.4%** lys 25.9% Gln 40.5% **Glu 9.6%**	K			K			**N** **2/13** K 11/13	**N** **14/15** K 1/15	**N** **5/14** K 5/14 R 4/14	K 15/15	K 7/12 **E** **4/12** T 1/12	**E** **3/13** K 10/13	K 12/12	K 6/8 R 1/8 **E** **1/8**	K 5/14 Q 8/14 R 1/14	G 13/13	**E** **14/19** R 5/19	K 14/14	K 19/19	K 11/11	**E** **2/7** Q 2/10
346	Val 37.6% Ala 26.8% Ser 11.2% **Phe 0.2%** **Tyr 0%**	V			A 5/16 V 11/16	V 15/15	V 10/10	V 4/13 A 4/13	V 15/15	V 10/14 A 4/14	V			V 3/12 **F** **5/12** A 4/12	V 6/8 **F** **2/8**	V 6/14 **Y** **8/14**	V			V 19/19	V 7/11 A 4/11	A 9/9

Table 1: Evolution of positively selected amino acids that were rarely found in the 500 sequences database. *(Continued)*

394	Thr 81.3% Iso 3.1% Ala 1.7%	T	T	T 4/13	T 13/15	T 6/14	T	T	T	T
				I 9/13	I 1/15 A 1/15	I 8/14				
476	Arg 81.3% Lys 18.4% Met 0%	K	R	R 11/13	R 3/15	R 8/14	R	**M**	R	R
				K 2/13	K 12/15	K 6/14				

Their frequency in the sequence database and their proportion (number of clones with the mutation/number of clones sequenced) in the viral quasispecies at each time point (I, II, and III) are shown.

shield that protects the virus against host defences [57]. Interestingly, it has been demonstrated that the neutralizing activity of a human monoclonal antibody, designated as mAb 2G12, is associated with the presence of glycosylation sites at these positions, including 295, 332 and 339 [58-60]. IgG 2G12-like antibodies have previously been detected in LTNP patients by competitive ELISA experiments with high levels in sera associated with the broad neutralizing activity [19]. This observation is in perfect agreement with our data, suggesting that antibodies that bind close to the 2G12 binding site exist in some patients and exert selective pressure on the viral surface.

It has recently been observed that cross-neutralizing activity characterizing a small subset of LTNPs is associated with antibodies recognizing the CD4bs [22]. However, only a few broadly neutralizing human monoclonal antibodies have been isolated at present; among them, only the IgGb12 (directed against the CD-4bs) and mAb 2G12 (recognizing oligomannose residues) target the gp120 [58,61,62]. Notably, 4 out of the 5 LTNP patients exhibit strong selective constraints at the level of the CD4bs. In patient F in particular, an IgGb12 epitope-like area is under strong positive selection (Figure 5). These data document that this epitope can be modified in vivo in response to specific selective pressure. Further analyses are necessary to clarify if mutations in this region may alter the viral RC, thus being able to delay disease progression.

Conclusion
The present study describes the dynamic evolution of the HIV-1 *env* gene in a subset of LTNP subjects and documents that the CD4bs is under strong selective pressure in the replicating virus of a group of LTNPs and evolves during the course of the disease. These data may be of interest not only for the understanding of the complex HIV-1-host relationships, but also for planning new immune-based strategies against HIV-1 infection.

Methods
Patients and sequences
Seven HIV-positive patients, never treated for HIV infection, were selected on the basis of the slope of their CD4+-T-cell counts and the level of HIV-1 viremia (Table 2). Two of them (subjects A and B) were showing a typical progression of HIV infection (TP), with a gradual decline of CD4+ T cells over time (loss of circulating CD4+ T cells per year: subject A, 87; and subject B, 153). The patients designated C, D, E, F, and G were LTNPs with CD4+-T-cell constantly higher than 500 per ml. They were showing the following mean variation/year of circulating CD4+lymphocytes: -31, -24, -2, -10, and +12 respectively). Lengths of infection and sampling dates are shown in Table 2.

Plasma specimens were concentrated by centrifugation at $23,600 \times g$ for 1 h at 4°C and RNA was extracted by using a QIAamp viral RNA mini kit (Qiagen, Valencia, CA). The following outer primers were used in the nested PCR amplification reaction: V31 (nucleotides 6939 to 6966 in the pNL4-3 numbering system) and V52 (nucleotides 7803 to 7778). The internal primers were: V32 (nucleotides 7367 to 7340), and V41 (nucleotides 7304 to 7326). The reverse transcription of HIV-1 RNA present in plasma was performed with primer V52 (25 pmol) and 200 U of SuperScript II RNase H-RT (Bethesda Research Laboratories, Gaithersburg, Md.) at 37°C for 60 min in a final volume of 20 µl in the presence of 3.0 mM $MgCl_2$, 75 mM KCl, 50 mM Tris (pH 8.3), 10 mM dithiothreitol, 0.5 mM concentrations of each deoxynucleosidetriphosphate (dNTP), and 20 U of recombinant RNasin RNase inhibitor (Promega Corp., Madison, Wis.). An amount of cDNA equivalent to 50–1000 copies of template (as evaluated by HIV-1 RNA copies/ml) was used for PCR amplification. An Expand High Fidelity PCR system (Roche Diagnostic Corporation, Indianapolis, IN) in 1× Expand PCR buffer containing 1.5 mM $MgCl_2$, 0.2 mM of each deoxynucleoside triphosphate, and 0.2 µM of V31 and V52 primers was used for the first round. The following cycling conditions were applied: 94°C for 2 min followed by 35 cycles of 94°C for 15 s, 55°C for 30 s, and 68°C for 4 min, with a final extension of 68°C for 10 min. Second-round PCR was performed using 2 µl of the first-round PCR product and primers V32 and V41 under the same conditions used for the first-round PCR.

Only one sample at a time was processed, and clinical samples were amplified in triplicate. Before molecular cloning, a 10-µl aliquot of the amplified product was run on a 10% polyacrylamide gel electrophoresis to screen for the appropriate-sized band (ca. 865 bp); the remaining 90 µl was resolved by electrophoresis on a 1.5% low-melting-point agarose gel (SeaPlaque; FMC BioProducts, Rockland, Maine) in TAE buffer (Tris-acetate, 1 mM EDTA). The DNA fragment was excised from the gel, purified by the QIAquick DNA Clean-Up system (Qiagen GmbH, Hiden, Germany) and cloned into pGEM-T vector (Promega) according to the manufacturer's instructions. After PCR colony screening, 8 to 20 positive clones were selected and the insert was sequenced with primers V31 and V52 on an automatic sequencer (ABI Prism 3100, Appliedbiosystems, Foster City, CA).

For patient A, 15 clones per time point were studied; for patient B 16,15 and 10 clones respectively; for patient C, 13, 15, and 24 clones; for patient D, 15,12 and 13; for patient E, 12, 8, and 14 clones, for patient F, 13, 19, and 14 clones; and for patient G, 19, 15, and 9 clones.

Table 2: Immunologic and virologic parameters of patients that were selected for the study.

Patient	Timepoint (years from infection)	HIVRNA copies/ml of plasma	CD4+ T cell counts/mm^3
A	I (9)	28000	561
	II(11)	32360	379
	III(13)	127200	337
B	I (8.5)	25000	374
	II(9.5)	65000	255
	III(10.5)	40000	259
C	I(10)	99	1144
	II(11)	100	944
	III(12)	549	776
D	I(9.5)	2918	993
	II(10.5)	3318	735
	III(12)	6850	640
E	I(11.5)	1200	795
	II(12.5)	2000	720
	III(13.2)	3034	527
F	I(8)	2100	645
	II(10)	1200	657
	III(13)	6393	625
G	I(10)	560	864
	II(13)	106	926
	III(14)	152	750

Because some analyses could potentially be affected by recombinats, we tested for recombination using the PHI test implemented in the SplitsTree package version 4.8. Significance of the PHI statistic for the presence of recombination is assessed with the normal approximation of a permutation test where, under the null hypothesis of no recombination, sites along the alignment are randomly permuted to obtain the null distribution of PHI: $p < 0.05$ indicate significant presence of recombination. About 15% of sequences for each time point/patient were discarded after this analysis.

GenBank accession numbers of sequences are EU329847 – EU330175.

High resolution phylogenetic analysis & graphical 3-dimensional visualization
Overall, 278 non recombinant genomic-HIV-1 RNA sequences were aligned collectively and individually for each patient, using amino acidic sequences as template for nucleotide alignment by using DAMBE http://dambe.bio.uottawa.ca/software.asp and manually corrected with BioEdit http://www.mbio.ncsu.edu/BioEdit/

bioedit.html. A Neighbor-Joining (NJ) tree was reconstructed using the best fitting evolutionary model as evaluated with MODELTEST v3.6 (Posada, D., and K.A. Crandall, 2001) was generated.

To obtain a maximum-likelihood tree topology, a local rearrangement search with the maximum-likelihood method was conducted by starting from the topology of the NJ tree, as implemented in PAUP* http://paup.csit.fsu.edu. The ratio of transitions to transversions, and the gamma distribution of rate variation among sites were estimated from the data. To evaluate if intra-patient virus evolution showed patterns of positive selection, a ML method was applied by using CODEML implemented in the PAML package http://apt.bea.ki.se/packages.html and the substitution rate at individual codon position was also estimated for each patient using the TipDate model as implemented in BEAST http://evolve.zoo.ox.ac.uk/Beast. The CODEML program fits various models of codon evolution to sequence data related by a phylogenetic tree, which allow to test for varying selection pressures at individual codon sites. The models of codon evolution differ in their distribution of dN/dS values among codons. Two couples of nested models were employed: M1a vs M2a and M7 vs M8. M1a (neutral/purifying model) allows only two categories of dN/dS across codons and the dN/dS ratio is constrained to be > 0 and < 1 in one category and equal to 1 in the other. Hence, M1a only accommodates neutral evolution. M2a adds an extra class of codons to account for positive selection (i.e., a class of codons with dN/dS > 1). M7 (neutral model) assumes a beta distribution of dN/dS between 0 and 1 with 10 categories to discretize the distribution. M8 adds an extra class of codons with dN/dS > 1 [63]. The likelihoods of the models were than compared using the likelihood ratio test. To allow further definition of HIV-1 env positively selected sites within each patient, all the RNA-sequences amplified and cloned from the same subject, were analysed by using an empirical Bayes approach [64]. The posterior mean dN/dS value per site was calculated and a Bayesian approach was used to identify codons undergoing positive selection with a posterior probability of > 95% or > 99%, using CODEML. To better identify conformational epitopes and sites on the protein surface with possibly distinct roles on disease progression, and distinct patterns of virus evolution driven by host-selective constraints along the C2-V5 region, a graphic colour-grade 3-dimensional visualization the dN/dS score (ratio between non-synonymous/synonymous mutations per site) was generated using PyMol http://www.pymol.sourceforge.net and the structure of a V3-containing gp120 core [65].

Moreover, to better understand the impact of deleterious mutational load in within-host HIV evolution and its impact on identifying positively selected sites, we performed a ML analysis of varying selection pressures among lineages. In this analysis, we compared model M0 (all branches have the same dN/dS) with an alternative model that allows a different dN/dS for internal and external branches.

Sites under selection along internal branches of reconstructed phylogenetic trees, were inferred using the Fixed Effects Likelihood (FEL) approach implemented in Hy-Phy http://www.hyphy.org.

To evaluate the HIV-1 general site-specific inter-exchangeability (site-specific aminoacidic entropy) a collection of 500 aligned env sequences from the Stanford Database was downloaded and analysed with BioEdit accessory applications. When positional homology was not maintained due to the high genetic variability, that site in the alignment was not considered in the analyses.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
FC conceived and coordinated the study and wrote the manuscript; FC and PL did the analyses. MCM, MS, SB and PB carried out the amplification and sequencing. MD did the 3D color-grade mapping on the gp120 structure. GG followed up patients; BM performed the HLA typing; RB, MC discussed the data and reviewed the manuscript. All authors read and approved the final manuscript.

Additional material

Additional file 1
Supplementary Table One. Likelihood ratio statistics (2Δl) for comparision of different models of codon evolution.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1742-4690-6-4-S1.doc]

Additional file 2
Supplementary Table Two. Analysis of selective pressure among internal and external branches in each patient.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1742-4690-6-4-S2.doc]

Additional file 3
Supplementary Table Three. Low- or high-resolution HLA typing.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1742-4690-6-4-S3.doc]

Acknowledgements

References

1. Richman DD, Little SJ, Smith DM, Wrin T, Petropoulos C, Wong JK: **HIV Evolution and Escape.** *Trans Am Clin Climatol Assoc* 2004, **115:**289-303.
2. Frost SD, Wrin T, Smith DM, Kosakovsky Pond SL, Liu Y, Paxinos E, Chappey C, Galovich J, Beauchaine J, Petropoulos CJ, *et al.*: **Neutralizing antibody responses drive the evolution of human immunodeficiency virus type 1 envelope during recent HIV infection.** *Proc Natl Acad Sci USA* 2005, **102(51):**18514-18519.
3. Clementi M, Canducci F, Bagnarelli P, Menzo S: **Intra-host evolution of human immunodeficiency virus type 1 and viral fitness.** *New Microbiol* 2004, **27(2 Suppl 1):**41-44.
4. Quinones-Mateu ME, Ball SC, Marozsan AJ, Torre VS, Albright JL, Vanham G, van Der Groen G, Colebunders RL, Arts EJ: **A dual infection/competition assay shows a correlation between ex vivo human immunodeficiency virus type 1 fitness and disease progression.** *J Virol* 2000, **74(19):**9222-9233.
5. Arts EJ, Quinones-Mateu ME: **Sorting out the complexities of HIV-1 fitness.** *Aids* 2003, **17(5):**780-781.
6. Rangel HR, Weber J, Chakraborty B, Gutierrez A, Marotta ML, Mirza M, Kiser P, Martinez MA, Este JA, Quinones-Mateu ME: **Role of the human immunodeficiency virus type 1 envelope gene in viral fitness.** *J Virol* 2003, **77(16):**9069-9073.
7. Bagnarelli P, Vecchi M, Burighel N, Bellanova D, Menzo S, Clementi M, De Rossi A: **Genotypic and phenotypic correlates of the HIV Type 1 env gene evolution in infected children with discordant response to antiretroviral therapy.** *AIDS Res Hum Retroviruses* 2004, **20(12):**1306-1313.
8. Menzo S, Sampaolesi R, Vicenzi E, Santagostino E, Liuzzi G, Chirianni A, Piazza M, Cohen OJ, Bagnarelli P, Clementi M: **Rare mutations in a domain crucial for V3-loop structure prevail in replicating HIV from long-term non-progressors.** *Aids* 1998, **12(9):**985-997.
9. Frost SD, Liu Y, Pond SL, Chappey C, Wrin T, Petropoulos CJ, Little SJ, Richman DD: **Characterization of human immunodeficiency virus type 1 (HIV-1) envelope variation and neutralizing antibody responses during transmission of HIV-1 subtype B.** *J Virol* 2005, **79(10):**6523-6527.
10. Riddle TM, Shire NJ, Sherman MS, Franco KF, Sheppard HW, Nelson JA: **Sequential turnover of human immunodeficiency virus type 1 env throughout the course of infection.** *J Virol* 2006, **80(21):**10591-10599.
11. Srivastava IK, Ulmer JB, Barnett SW: **Role of neutralizing antibodies in protective immunity against HIV.** *Hum Vaccin* 2005, **1(2):**45-60.
12. Moore PL, Crooks ET, Porter L, Zhu P, Cayanan CS, Grise H, Corcoran P, Zwick MB, Franti M, Morris L, *et al.*: **Nature of nonfunctional envelope proteins on the surface of human immunodeficiency virus type 1.** *J Virol* 2006, **80(5):**2515-2528.
13. Lee WR, Syu WJ, Du B, Matsuda M, Tan S, Wolf A, Essex M, Lee TH: **Nonrandom distribution of gp120 N-linked glycosylation sites important for infectivity of human immunodeficiency virus type 1.** *Proc Natl Acad Sci USA* 1992, **89(6):**2213-2217.
14. Wei X, Decker JM, Wang S, Hui H, Kappes JC, Wu X, Salazar-Gonzalez JF, Salazar MG, Kilby JM, Saag MS, *et al.*: **Antibody neutralization and escape by HIV-1.** *Nature* 2003, **422(6929):**307-312.
15. Balzarini J: **Targeting the glycans of gp120: a novel approach aimed at the Achilles heel of HIV.** *Lancet Infect Dis* 2005, **5(11):**726-731.
16. Poon AF, Lewis FI, Pond SL, Frost SD: **Evolutionary interactions between N-linked glycosylation sites in the HIV-1 envelope.** *PLoS Comput Biol* 2007, **3(1):**e11.
17. Burton DR, Desrosiers RC, Doms RW, Koff WC, Kwong PD, Moore JP, Nabel GJ, Sodroski J, Wilson IA, Wyatt RT: **HIV vaccine design and the neutralizing antibody problem.** *Nat Immunol* 2004, **5(3):**233-236.
18. Zolla-Pazner S: **Identifying epitopes of HIV-1 that induce protective antibodies.** *Nat Rev Immunol* 2004, **4(3):**199-210.
19. Braibant M, Brunet S, Costagliola D, Rouzioux C, Agut H, Katinger H, Autran B, Barin F: **Antibodies to conserved epitopes of the HIV-1 envelope in sera from long-term non-progressors: prevalence and association with neutralizing activity.** *Aids* 2006, **20(15):**1923-1930.
20. Eggink D, Melchers M, Sanders RW: **Antibodies to HIV-1: aiming at the right target.** *Trends Microbiol* 2007, **15(7):**291-294.
21. Dhillon AK, Donners H, Pantophlet R, Johnson WE, Decker JM, Shaw GM, Lee FH, Richman DD, Doms RW, Vanham G, *et al.*: **Dissecting the neutralizing antibody specificities of broadly neutralizing sera from human immunodeficiency virus type 1-infected donors.** *J Virol* 2007, **81(12):**6548-6562.
22. Li Y, Migueles SA, Welcher B, Svehla K, Phogat A, Louder MK, Wu X, Shaw GM, Connors M, Wyatt RT, *et al.*: **Broad HIV-1 neutralization mediated by CD4-binding site antibodies.** *Nat Med* 2007, **13(9):**1032-1034.
23. Moore PL, Gray ES, Choge IA, Ranchobe N, Mlisana K, Abdool Karim SS, Williamson C, Morris L: **The c3-v4 region is a major target of autologous neutralizing antibodies in human immunodeficiency virus type 1 subtype C infection.** *J Virol* 2008, **82(4):**1860-1869.
24. Kupfer B, Sing T, Schuffler P, Hall R, Kurz R, McKeown A, Schneweis KE, Eberl W, Oldenburg J, Brackmann HH, *et al.*: **Fifteen years of env C2V3C3 evolution in six individuals infected clonally with human immunodeficiency virus type 1.** *J Med Virol* 2007, **79(11):**1629-1639.
25. Bunnik EM, Pisas L, van Nuenen AC, Schuitemaker H: **Autologous neutralizing humoral immunity and evolution of the viral envelope in the course of subtype B human immunodeficiency virus type 1 infection.** *J Virol* 2008, **82(16):**7932-7941.
26. Joos B, Fischer M, Schweizer A, Kuster H, Boni J, Wong JK, Weber R, Trkola A, Gunthard HF: **Positive in vivo selection of the HIV-1 envelope protein gp120 occurs at surface-exposed regions.** *J Infect Dis* 2007, **196(2):**313-320.
27. Choisy M, Woelk CH, Guegan JF, Robertson DL: **Comparative study of adaptive molecular evolution in different human immunodeficiency virus groups and subtypes.** *J Virol* 2004, **78(4):**1962-1970.
28. Burton DR, Stanfield RL, Wilson IA: **Antibody vs. HIV in a clash of evolutionary titans.** *Proc Natl Acad Sci USA* 2005, **102(42):**14943-14948.
29. Gaschen B, Taylor J, Yusim K, Foley B, Gao F, Lang D, Novitsky V, Haynes B, Hahn BH, Bhattacharya T, *et al.*: **Diversity considerations in HIV-1 vaccine selection.** *Science* 2002, **296(5577):**2354-2360.
30. Scanlan CN, Offer J, Zitzmann N, Dwek RA: **Exploiting the defensive sugars of HIV-1 for drug and vaccine design.** *Nature* 2007, **446(7139):**1038-1045.
31. Leal E, Janini M, Diaz RS: **Selective pressures of human immunodeficiency virus type 1 (HIV-1) during pediatric infection.** *Infect Genet Evol* 2007, **7(6):**694-707.
32. de Oliveira T, Salemi M, Gordon M, Vandamme AM, van Rensburg EJ, Engelbrecht S, Coovadia HM, Cassol S: **Mapping sites of positive selection and amino acid diversification in the HIV genome: an alternative approach to vaccine design?** *Genetics* 2004, **167(3):**1047-1058.
33. Pybus OG, Rambaut A, Belshaw R, Freckleton RP, Drummond AJ, Holmes EC: **Phylogenetic evidence for deleterious mutation load in RNA viruses and its contribution to viral evolution.** *Mol Biol Evol* 2007, **24(3):**845-852.
34. Zhou T, Xu L, Dey B, Hessell AJ, Van Ryk D, Xiang SH, Yang X, Zhang MY, Zwick MB, Arthos J, *et al.*: **Structural definition of a conserved neutralization epitope on HIV-1 gp120.** *Nature* 2007, **445(7129):**732-737.
35. Salemi M, Burkhardt BR, Gray RR, Ghaffari G, Sleasman JW, Goodenow MM: **Phylodynamics of HIV-1 in Lymphoid and Non-Lymphoid Tissues Reveals a Central Role for the Thymus in Emergence of CXCR4-Using Quasispecies.** *PLoS ONE* 2007, **2(9):**e950.
36. Pond SL, Frost SD, Grossman Z, Gravenor MB, Richman DD, Brown AJ: **Adaptation to different human populations by HIV-1 revealed by codon-based analyses.** *PLoS Comput Biol* 2006, **2(6):**e62.

37. Yamaguchi-Kabata Y, Gojobori T: **Reevaluation of amino acid variability of the human immunodeficiency virus type 1 gp120 envelope glycoprotein and prediction of new discontinuous epitopes.** *J Virol* 2000, **74(9):**4335-4350.
38. Sheridan I, Pybus OG, Holmes EC, Klenerman P: **High-resolution phylogenetic analysis of hepatitis C virus adaptation and its relationship to disease progression.** *J Virol* 2004, **78(7):**3447-3454.
39. Bazykin GA, Dushoff J, Levin SA, Kondrashov AS: **Bursts of nonsynonymous substitutions in HIV-1 evolution reveal instances of positive selection at conservative protein sites.** *Proc Natl Acad Sci USA* 2006, **103(51):**19396-19401.
40. Lemey P, Rambaut A, Pybus OG: **HIV evolutionary dynamics within and among hosts.** *AIDS Rev* 2006, **8(3):**125-140.
41. Lemey P, Kosakovsky Pond SL, Drummond AJ, Pybus OG, Shapiro B, Barroso H, Taveira N, Rambaut A: **Synonymous substitution rates predict HIV disease progression as a result of underlying replication dynamics.** *PLoS Comput Biol* 2007, **3(2):**e29.
42. McMichael A, Klenerman P: **HIV/AIDS. HLA leaves its footprints on HIV.** *Science* 2002, **296(5572):**1410-1411.
43. Carrington M, O'Brien SJ: **The influence of HLA genotype on AIDS.** *Annu Rev Med* 2003, **54:**535-551.
44. Peyerl FW, Bazick HS, Newberg MH, Barouch DH, Sodroski J, Letvin NL: **Fitness costs limit viral escape from cytotoxic T lymphocytes at a structurally constrained epitope.** *J Virol* 2004, **78(24):**13901-13910.
45. Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, Zhang K, Gumbs C, Castagna A, Cossarizza A, *et al.*: **A whole-genome association study of major determinants for host control of HIV-1.** *Science* 2007, **317(5840):**944-947.
46. Moore CB, John M, James IR, Christiansen FT, Witt CS, Mallal SA: **Evidence of HIV-1 adaptation to HLA-restricted immune responses at a population level.** *Science* 2002, **296(5572):**1439-1443.
47. Mallal S, Nolan D, Witt C, Masel G, Martin AM, Moore C, Sayer D, Castley A, Mamotte C, Maxwell D, *et al.*: **Association between presence of HLA-B* HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir.** *Lancet* 5701, **359(9308):**727-732.
48. Bailey JR, Zhang H, Wegweiser BW, Yang HC, Herrera L, Ahonkhai A, Williams TM, Siliciano RF, Blankson JN: **Evolution of HIV-1 in an HLA-B*57-positive patient during virologic escape.** *J Infect Dis* 2007, **196(1):**50-55.
49. Dong T, Stewart-Jones G, Chen N, Easterbrook P, Xu X, Papagno L, Appay V, Weekes M, Conlon C, Spina C, *et al.*: **HIV-specific cytotoxic T cells from long-term survivors select a unique T cell receptor.** *J Exp Med* 2004, **200(12):**1547-1557.
50. Wang S, Sun Y, Zhai S, Zhuang Y, Zhao S, Kang W, Li X, Huang D, Yu XG, Walker BD, *et al.*: **Identification of HLA-A11-restricted HIV-1-specific cytotoxic T-lymphocyte epitopes in China.** *Curr HIV Res* 2007, **5(1):**119-128.
51. Gnanakaran S, Lang D, Daniels M, Bhattacharya T, Derdeyn CA, Korber B: **Clade-specific differences between human immunodeficiency virus type 1 clades B and C: diversity and correlations in C3-V4 regions of gp120.** *J Virol* 2007, **81(9):**4886-4891.
52. Pinter A: **Roles of HIV-1 Env variable regions in viral neutralization and vaccine development.** *Curr HIV Res* 2007, **5(6):**542-553.
53. Pantophlet R, Burton DR: **GP120: target for neutralizing HIV-1 antibodies.** *Annu Rev Immunol* 2006, **24:**739-769.
54. Vaine M, Wang S, Crooks ET, Jiang P, Montefiori DC, Binley J, Lu S: **Improved induction of antibodies against key neutralizing epitopes by human immunodeficiency virus type 1 gp120 DNA prime-protein boost vaccination compared to gp120 protein-only vaccination.** *J Virol* 2008, **82(15):**7369-7378.
55. Karlsson Hedestam GB, Fouchier RA, Phogat S, Burton DR, Sodroski J, Wyatt RT: **The challenges of eliciting neutralizing antibodies to HIV-1 and to influenza virus.** *Nat Rev Microbiol* 2008, **6(2):**143-155.
56. McKnight A, Aasa-Chapman MM: **Clade specific neutralising vaccines for HIV: an appropriate target?** *Curr HIV Res* 2007, **5(6):**554-560.
57. Dacheux L, Moreau A, Ataman-Onal Y, Biron F, Verrier B, Barin F: **Evolutionary dynamics of the glycan shield of the human immunodeficiency virus envelope during natural infection and implications for exposure of the 2G12 epitope.** *J Virol* 2004, **78(22):**12625-12637.
58. Mehandru S, Vcelar B, Wrin T, Stiegler G, Joos B, Mohri H, Boden D, Galovich J, Tenner-Racz K, Racz P, *et al.*: **Adjunctive passive immunotherapy in human immunodeficiency virus type 1-infected individuals treated with antiviral therapy during acute and early infection.** *J Virol* 2007, **81(20):**11016-11031.
59. Manrique A, Rusert P, Joos B, Fischer M, Kuster H, Leemann C, Niederost B, Weber R, Stiegler G, Katinger H, *et al.*: **In vivo and in vitro escape from neutralizing antibodies 2G12, 2F5, and 4E10.** *J Virol* 2007, **81(16):**8793-8808.
60. Scanlan CN, Pantophlet R, Wormald MR, Ollmann Saphire E, Stanfield R, Wilson IA, Katinger H, Dwek RA, Rudd PM, Burton DR: **The broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2G12 recognizes a cluster of alpha1-->2 mannose residues on the outer face of gp120.** *J Virol* 2002, **76(14):**7306-7321.
61. Trkola A, Kuster H, Rusert P, Joos B, Fischer M, Leemann C, Manrique A, Huber M, Rehr M, Oxenius A, *et al.*: **Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies.** *Nat Med* 2005, **11(6):**615-622.
62. Nakowitsch S, Quendler H, Fekete H, Kunert R, Katinger H, Stiegler G: **HIV-1 mutants escaping neutralization by the human antibodies 2F5, 2G12, and 4E10: in vitro experiments versus clinical studies.** *Aids* 2005, **19(17):**1957-1966.
63. Yang Z: **Maximum likelihood analysis of adaptive evolution in HIV-1 gp120 env gene.** *Pac Symp Biocomput* 2001:226-237.
64. Yang Z, Wong WS, Nielsen R: **Bayes empirical bayes inference of amino acid sites under positive selection.** *Mol Biol Evol* 2005, **22(4):**1107-1118.
65. Huang CC, Tang M, Zhang MY, Majeed S, Montabana E, Stanfield RL, Dimitrov DS, Korber B, Sodroski J, Wilson IA, *et al.*: **Structure of a V3-containing HIV-1 gp120 core.** *Science* 2005, **310(5750):**1025-1028.